REVIEW AND ANALYSIS ON NETWORK ON CHIP

G. Vishnu Vardhan Rao
.Assistant professor
Department of Electronics and Communication Engineering
Vel Tech Multi Tech Dr. Rangarajan Dr. Sakunthala Engineering College,
Chennai,India,
gadipudir@gmail.com

Dr. A. Kavitha,
professor
Department of Electronics and Communication Engineering
M. Kumarasamy College of Engineering,
Karur,India,
kavivenkat99@gmail.com

Arthy P S
Assistant professor,
Department of Electronics and Communication Engineering
Sri Sai Ram Institute of Technology
Chennai,India,
arthy.ece@sairamit.edu.in

2022 International Conference on Computer, Power and Communications (ICCPC) | 979-8-3503-9784-0/22/$31.00 ©2022 IEEE | DOI: 10.1109/ICCPC55978.2022.10072109

Abstract

Moore's law has been extended with the introduction of multiprocessor architectures and platforms. To boost design productivity and system performance, they leverage concurrency and synchronisation in both software and hardware. Furthermore, these systems must be highly scalable, reusable, and repeatable. They'll also have to be low-cost and low-energy. It is expected that deep sub-micron technology, defined by gate lengths of 60-90 nm, will be plagued by wire delays, signal integrity issues, and unsynchronized communications as billion transistors age. The important developments in NOC research are discussed, as well as aspects that demand more exploration. The most widely used and researched topology is a packet-switched 2-D mesh. It's also a bit of an average NoC right now. There are lots of positive outcomes and intriguing proposals. However, there were significant discrepancies in implementation results, ambiguous documentation, and a lack of comparability.

Keywords: Embedded Systems, Network-On-Chip (NOC), Literature study, SOC, challenges.

1. INTRODUCTION

A simple SOC consists of a CPU, instruction and data cache attached to the CPU, and numerous peripherals such as DMA, LCD, BIU Instruction bus, CAN bus, Instruction memory, data memory, and other peripherals. In addition to these slower peripherals, roughly 20 peripherals are conceptually integrated and communicate utilising the AMBA architecture, including WDT, Interrupt controller, timer, and real timeclocks. AMBA is a RISC 5 processor that works with ARM processors and other RISC 5 processors. However, today's SOC is more complicated, with many buses. Critical pathways and clock trees are commonly used in conventional system architectures. The increased power usage is due to these crucial paths and clock trees. As a result, SOCs are inefficient in terms of power consumption. Furthermore, due to time skew issues, it is difficult to manage these clock trees. This method eliminates the need for chip-wide clocktrees, allowing designers to concentrate on local synchronous zones, which are significantly less complex than the entire system. There is more control over how much power is consumed in a system when only one synchronous zone (or node) can have its clock speed altered without affecting the other synchronous zones. The NOC is a GAL's solution (Network-on-Chip). This trend highlights the need of employing a NoC solution from a company with a track record in the market to speed up the creation of new chips. Arteris is now the sole prominent player in the NoC scene; nevertheless, as demonstrated in Table 1, Thus, new avenues for innovation and the creation of products for new market niches are now accessible.

NOCs were purchased by large Silicon Valley firms. Qualcomm purchased Arteris in part, preserving the patents listed in table 1.

Functional, power, and economic needs can be met while operational safety and data dependability are ensured by a NoC when both are required. When creating semiconductors for today's applications, the connection networks within and outside the device must have sufficient capacity.

Table 1 displays some of the most important electronics companies' chip innovations. With tens of thousands of computing cores, these new developments pose major challenges to chip connections. An on-chip networking system is used by all companies to connect processing cores to each other. Using this example, it is clear that NOCs are essential to the industry.

2. SYSTEMATIC LITERATURE REVIEW

2017 Seema On a large scale the scalability of microchip technologies has enabled System-on-Chip (SoC). As data-intensive applications proliferated and computing power grew, the threat of communication components on single-chip systems arose, resulting in the introduction of network on chip (NoC). The concept of interachip communication is provided by NoC. In this paper, an excellent concept for system-on-chip communication known as communication network on-

Company	Buyer
Arteris(partial)	Qualcomm
Netspeed	Intel
Sonics	Facebook

Table 1

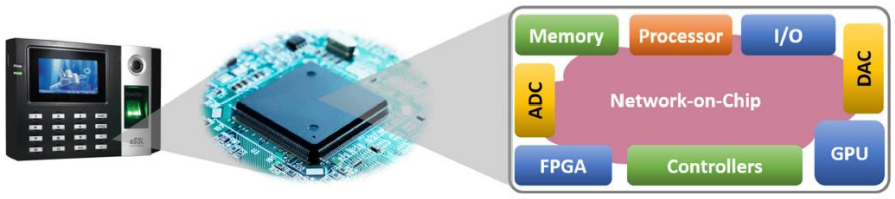

Fig. 1. Systems on Chip (SOCs) that use a NOC-based communication fabric and IP Processes are shown.

chip is discussed (NoC). The fundamentals of NoC, network topology, significant research topics, and various abstraction levels are all covered in this study.

For a SoC with a large number of processing units and memory, the interconnection of such pieces becomes increasingly critical in the design process. SoC architectures have seen an increase in the types and quantities of hardware accelerators with each successive chip generation. Hardware modification is required for algorithm acceleration as well as dataflow optimization in today's data processing approaches. In a Many-Core SoC, a NoC allows designers to segregate compute and communication issues. As a result, a flexible NoC is critical for system performance. Scalable networks with thousands of processing cores are not straightforward to design, and addressing different challenges in a modular fashion requires a layered approach. The administration of a network with tens of thousands of components is both challenging and fascinating due to the increasing complexity of the NoC architecture as a result of dynamic demand [2].

The SDN concept addresses the management issue in NoC-based Many-Core systems. A Benefits of Software-Defined Network on a Chip (SDNoC) include greater runtime flexibility and self-adaptive network management. It is possible to simplify NoC routers by utilising an SDNoC controller to provide features like Quality of Service (QoS), fault tolerance and power management. Because of this, routers can be customised, and their major function is to reroute NoC packets according to the SDNoC controller's policies and regulations. The SDNoC controller observes the NoC from a distance. As a result, the SDNoC controller optimises networks worldwide or many objectives, including fault mitigation, load balancing, power consumption, and quality of service for real-time flows. In addition, the controller can incorporate various objectives in order to achieve multi-objective management [3].

In the NoC-based design, several methodologies and techniques from the classic computer networking methodology are used [4]. For on-chip communication, the NoC idea uses a packet-switching fabric. As shown in Figure 2.1, the computational capacity of the SoC is divided into processing tiles, each of which is made up of processor elements, memory, Network Interfaces (NI), and routers. The compute and communication domains are separated by Network Interfaces. At the source, NI packetizes the data coming from the CPU cores. NI assembles the message from all of the packets and sends it to the destination's processing core. Routers and links are the most important NoC components. Networks that connect two routers are known as point-to-point networks, and routers send packets between the two nodes.

Due to reduced transistor geometries and strict voltage scaling, current technology allows us to combine hundreds of processing cores, but it also introduces some reliability difficulties. As demonstrated in Table 2.1, the susceptibility of created NoC components affects their dependability and is caused by transient and permanent failures. As a result, improving system reliability is a substantial task [5].

With a NoC, operational safety and data dependability can be ensured, as well as power and cost needs being met in a cost-effective manner. When creating semiconductors for today's applications, the connection networks within and outside the device must have sufficient capacity. Table 2.2 illustrates some of the most important electronics companies' chip developments. Some of these advancements have tens of thousands of computing cores, posing significant obstacles to chip interconnection [6,7]. It's worth noting that all organisations use an on-Chip network to link the processor units together. This data emphasises the industry's relevance of NoCs.

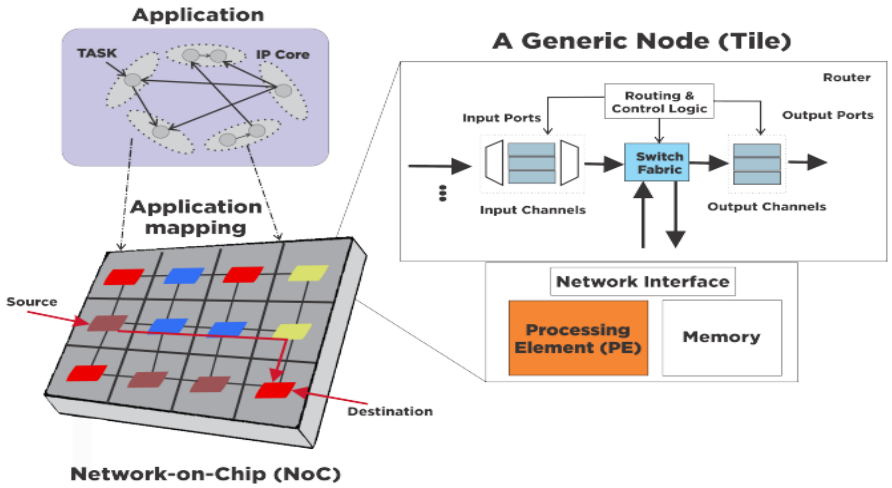

Figure 2.1. The NoC application, its architecture, and its relationship to the NoC. There is a node micro-architecture illustrated on the right-hand side of the picture that includes an on-chip router, buffers, and a processing element (PE).

Type of Fault	Syndrome	Characteristics	Cause	Solution
Transient	One or more bit-errors in a transmitted packet	Random and short duration	Crosstalk effect or alpha or neutron particle strikes	Error checking/ correcting schemes or retransmissions
Permanent	Faulty elements	Non-recoverable device defects	Negative Bias Temperature Instability (NBTI), Hot Carrier Injection (HCI), Time-Dependent Dielectric Breakdown (TDDB)	Redundant resources available to replace the faulty components

Table 2.1. NoC reliability is impacted by a variety of chip failures that exhibit different characteristics and have different sources.

Chip	Company	Features(Cores)	NoC	References	Year
Xeon Phi	Intel	72Xeoncores	Ring(Photonic) Knight's Ferry-Corner/2DMesh Knight's landing	[8]	2016
Open Piton+Ariane	-	65,536	2D Mesh crossbar	[9]	2019
DaVinci Ascend910	Huawei	32DaVincicores	Uses Arteris NoC Ring and Mesh bus	[10]	2019
TrueNorth	IBM	4096Neurosynapticcores	2DMesh	[11,12]	2019
CloudA100	Qualcomm	16Neuralprocessors	Arteris	[13,14]	2020
Loihi	Intel	128Neuromorphiccores,	A2D Mesh NoC with up to 4096 on-Chipcores.	[15]	2020
Epiphany-V	Adapteva	1024RISC-Vcores	Three NoCs	[16]	2020
AI Processor	-	16,384(128×128) Nanocores	7NoCs	[17]	2020

Table2.2. The most recent chips from the companies in question. We go over the different kinds of cores and how they are connected to one another on the chip. Many-Core architectures are the norm.

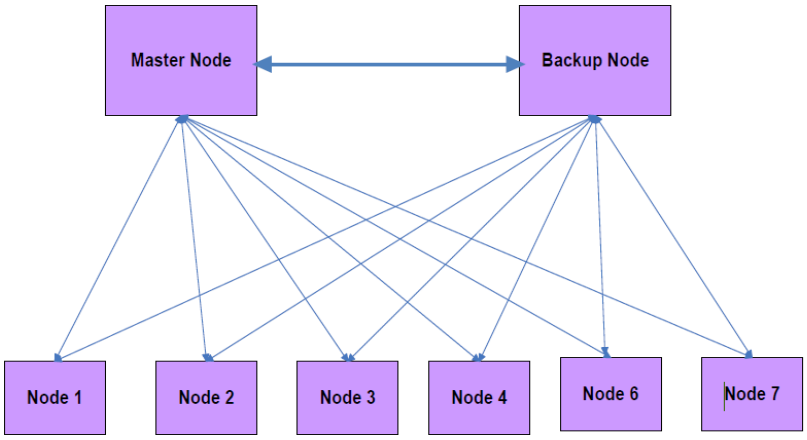

Figure.3.1. Dynamic adaption of the NoC while failure of master hardware modules.

3. METHODOLOGY

3.1 CURRENT CHALLENGES IN NOC DESIGN

The number of routes between cores grows by a squared ratio when more IP blocks are added to the chip, increasing the congestion problem. Congestion control has been one of the most difficult aspects of NOC design. Increases in the die's size and the number of metal mask layers have an effect on the timing closure, resulting in unanticipated paths. This makes developing low-cost, high-performance chips a big problem. Also, as technology advances, the driving intensity of the transistor decreases, but signal propagation time along the wire increases, affecting chip speed. The cost of the chip rises as the number of wires grows. Aside from that, the NOC designer must deal with challenges such as Complex Multi-Variable Problems, Large Design Space, Comfort around Simplistic Designs, Performance, Power and Area Tradeoffs, Buffer Sizing, and Pipelines. Because interconnect resources are limited and often prove to be a performance bottleneck, NoC should enable network-level congestion control as well. Buffering issues, channel width issues, and other issues are becoming more prevalent. NOC circuits should be built to overcome these challenges, as well as to reduce latency and enhance bandwidth.

Band width = fch*W

The channel's operational frequency is fch. To reduce latency, the band width must be raised. Total packet completion *Packet length/(Number of IP blocks)*time Equals packet throughput Apart from these, selecting the topology and routing algorithms are two more issues for NOC designers.

3.2 Multidrop links

There are links in this architecture where a signal reaches many routers in the same clock cycle.The diameter of the network gets is lowered in this architecture, which eliminates the requirement for high radix routers and thereby improves network latency.

By sending bypass signals towards the destination, these linkages can be formed and broken.

3.3 Three Dimensional NOC

The number of cores inside the chip is expanding at an exponential pace, lowering the performance of 2D ICs. There are also constraints on floor layout options, and 2D ICs have bigger die sizes in multiprocessor based applications. There are also many issues with clock distributions. The 3D architectures were born as a result of this. In a 3DNOC, a vertical interconnection of several 2Dlayers one above the other with Near field coupling schemes where there is no physical connection between the layers is communicated to upper levels without physical connections by inductive and capacitive coupling. 3D NOC is mostly utilised in real-time and aerospace applications to improve system performance. Designing a 3D IC presents a number of obstacles. The main benefits of using a wireless NOC include deadlock-free routing, low latency, multicasting, and broadcasting, among others.

3.4 System model

To cover a suitable design area, the modelling technique used must enable a variety of system sizes and configurations. Simulating and analysing all of the minute nuances is almost impossible, which inevitably leads to estimating inaccuracies. As a result, the entire simulation must be tested using a real-world implementation as well as analytical models or protocol specifications (designed protocol, traffic, environment model, and usage scenario). Though it may be less precise in the beginning, it can be improved upon later. According to the environment and the NoC models (transactional, cycle-accurate, area, and wire estimates), this (traffic

generation, third-party IP). It's risky to compare only the in-house simulation models with the analytical models because they both could be based on incorrect assumptions. As a result, validation should be done against external, independent data wherever possible.

3.5 Dynamically Reconfigurable Packet-Switched NoC

Packet switching and NoC Design modules make up the majority of the Dynamically Reconfigurable module. First, the receiver module receives data serially, similar to a FIFO, and sends it to RAM. The data is stored in address locations in the RAM. If the data is genuine, the analyzer examines it and sends it to the NoC in the form of a packet.

Routers, hardware modules, and I/O buffers are the core components of the NOC. As demonstrated in Figure 3.1, apply the dynamic adaptation idea to NOC.

When a NoC's structure fails, great care must be taken to avoid isolating parts of the NoC and to ensure that the collapse does not disrupt communication between the nodes. Figure 3 depicts a typical situation. A hardware module (Master Module) is claimed to have failed and is no longer capable of switching data between nodes, necessitating its replacement by a new one that fits in the same space. Because a connection port for the master node to the NoC is lacking in the new configuration, a new module must be introduced to restore communication without disrupting current communication. Because the dynamic reconfiguration must not affect other hardware modules or their connectivity, no packets must be transferred directly from any node to the recipient node via the master module. As a result, switch node routing tables are temporarily modified in order for communication to be established with the help of the Backup node. Once the new master node is configured, the nodes are updated with the new master node address so that they can continue to function normally.

CONCLUSION

Many-Core SoC NoC management difficulties necessitate a different approach than with regular NoCs because of their complexity. The layered SDNoC architecture is a modular method that allows for innovation at various abstraction layers due to the separation of responsibilities. SDNoC documentation is excellent, yet several SDN features are still missing. If management solutions are not reusable, it means that new designs and delays to market are more likely. Therefore, future SoC designs will not make advantage of an SDNoC architecture. Researchers may focus on these challenges without having to deal with the low-level specifics in an SDNoC architecture because NoC management comprises difficult optimization difficulties. We've also seen that a framework for optimising multi-core systems is essential.

REFERENCES

[1] Seema, “Network-on-Chip: A State-of-the-art Review” 2017.

[2] Ruaro, M.; Caimi, L.L.; Moraes, F.G. SDN-based Secure Application Admission and Execution for Many-cores. IEEE Access 2020, 8.

[3] Ruaro, M.; Caimi, L.L.; Moraes, F.G. A Systemic and Secure SDN Framework for NoC-Based Many-Cores. IEEE Access 2020, 8, 105997–106008.

[4] Benini, L.; De Micheli, G. Networks on chips: A new SoC paradigm. Computer 2002, 35, 70–78. [CrossRef]

[5] Chatterjee, N.; Mukherjee, P.; Chattopadhyay, S. Reliability-aware application mapping onto mesh based Network-on-Chip. Integration 2018, 62, 92–113.

[6] Balkind, J.; Lim, K.; Gao, F.; Tu, J.; Wentzlaff, D.; Schaffner, M.; Zaruba, F.; Benini, L. OpenPiton + Ariane: The First Open-Source, SMP Linux-booting RISC-V System Scaling From One to Many Cores. In Proceedings of the Third Workshop on Computer Architecture Research with RISC-V, CARRV, Phoenix, AZ, USA, 22 June 2019.

[7] Olofsson, A. Epiphany-v: A 1024 processor 64-bit risc system-on-chip. arXiv 2016, arXiv:1610.01832

[8] Jeffers, J.; Reinders, J.; Sodani, A. Knights Landing overview. In Intel Xeon Phi Processor High Performance Programming, 2nd ed.; Elsevier Inc.: Amsterdam, The Netherlands, 2016; pp. 15–24. [CrossRef]

[9] Liao, H.; Tu, J.; Xia, J.; Zhou, X. DaVinci: A Scalable Architecture for Neural Network Computing. In Proceedings of the 2019 IEEE Hot Chips 31 Symposium (HCS 2019), Cupertino, CA, USA, 18–20 August 2019. [CrossRef]

[10] Hsu, J. IBM’s new brain [News]. IEEE Spectr. 2014, 51, 17–19. [CrossRef]

[11] Debole, M.V.; Taba, B.; Amir, A.; Akopyan, F.; Andreopoulos, A.; Risk, W.P.; Kusnitz, J.; Otero, C.O.; Nayak, T.K.; Appuswamy, R.; et al. TrueNorth: Accelerating From Zero to 64 Million Neurons in 10 Years. Computer 2019, 52, 20–29. [CrossRef]

[12] Qualcomm. Cloud AI 100. 2020. Available online: https://www.qualcomm.com/products/cloud-artificial-intelligence/cloud-ai (accessed on 12 December 2020).

[13] Davies, M.; Srinivasa, N.; Lin, T.H.; Chinya, G.; Cao, Y.; Choday, S.H.; Dimou, G.; Joshi, P.; Imam, N.; Jain, S.; et al. Loihi: A Neuromorphic Manycore Processor with On-Chip Learning. IEEE Micro 2018, 38, 82–99. [CrossRef]

[14] Han, J.; Choi, M.; Kwon, Y. 40-TFLOPS artificial intelligence processor with function-safe programmable many-cores for ISO26262 ASIL-D. ETRI J. 2020, 42, 468–479. [CrossRef]

[15] Swarbrick, I.; Gaitonde, D.; Ahmad, S.; Gaide, B.; Arbel, Y. Network-on-chip programmable platform in versal™ ACAP architecture. In Proceedings of the 2019 ACM/SIGDA International Symposium on Field-Programmable Gate Arrays (FPGA 2019), Seaside, CA, USA, 24–26 February 2019; pp. 212–221. [CrossRef]

[16] Swarbrick, I.; Gaitonde, D.; Ahmad, S.; Jayadev, B.; Cuppett, J.; Morshed, A.; Gaide, B.; Arbel, Y. Versal network-on-chip (NoC). In Proceedings of the 2019 IEEE Symposium on High-Performance Interconnects (HOTI 2019), Santa Clara, CA, USA, 14–16 August 2019; pp. 13–17. [CrossRef]

[17] Xilinx. Versal ACAP Programmable Network on Chip and Integrated Memory; Technical Report; Xilinx: San Jose, CA, USA, 2020